AD-A086 661 MICHIGAN UNIV ANN ARBOR DEPT OF MATHEMATICS F/G 12/1
APPROXIMATION OF THE SPECTRUM OF CLOSED OPERATORS. THE DETERMIN--ETC(U)
MAR 80 M LUSKIN, J DESCLOUX, J RAPPAZ F49620-79-C-0149
UNCLASSIFIED AFOSR-TR-80-0548 NL

1 OF 1
AD
A086661

END
DATE
FILMED
8-80
DTIC

submitted to Mathematics of Computation

Dec. 21, 1979

(12)

APPROXIMATION OF THE SPECTRUM OF CLOSED OPERATORS
THE DETERMINATION OF NORMAL MODES OF A ROTATING BASIN

Jean Descloux
Département de Mathématiques
École Polytechnique Fédérale
1007 Lausanne
Switzerland

Mitchell Luskin[1]
Department of Mathematics
The University of Michigan
Ann Arbor, Michigan 48109

Jacques Rappaz[2]
Centre de Mathématiques Appliquées
École Polytechnique
91128 Palaiseau
France

[1]Supported by AFOSR under Contract F49620-79-C-0149 and by a Faculty Research Fellowship from the Horace Rackham School of Graduate Studies, The University of Michigan.

[2]Supported by the Fonds National Suisse de la Recherche Scientifique.

AMS(MOS) subject classifications (1970). Primary 65N25, 65N30.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(12)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (18) AFOSR-TR- 80-0548 (19)	2. GOVT ACCESSION NO. AD-A086 661
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) (6) APPROXIMATION OF THE SPECTRUM OF CLOSED OPERATORS, THE DETERMINATION OF NORMAL MODES OF A ROTATING BASIN	5. TYPE OF REPORT & PERIOD COVERED (9) Interim
	6. ~~PERFORMING ORG. REPORT NUMBER~~
7. ~~AUTHOR(s)~~ (10) Mitchell/Luskin, Jean/Descloux [redacted] Jacques/Rappaz	8. CONTRACT OR GRANT NUMBER(s) (15) F49620-79-C-0149
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Michigan, Department of Mathematics, Ann Arbor, MI 48109 (17) A3	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (16) 61102F 2304/A3
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332	12. REPORT DATE March 1980
	13. NUMBER OF PAGES 39
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) (12) 47 (11) Mar 80	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Submitted to MATHEMATICS OF COMPUTATION, December 1979

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper gives a theory of spectral approximation for closed operators in Banach spaces. The perturbation theory developed in this paper is applied to the study of a finite element procedure for approximating the spectral properties of a differential system modeling a fluid in a rotating basin.

Abstract

This paper gives a theory of spectral approximation for closed operators in Banach spaces. The perturbation theory developed in this paper is applied to the study of a finite element procedure for approximating the spectral properties of a differential system modeling a fluid in a rotating basin.

Introduction

In this paper, we give a theory of spectral approximation for closed operators in Banach spaces. We then apply this theory to an analysis of the approximation of the spectral properties of some differential systems by finite element methods.

Bramble and Osborn [1] and Osborn [14] developed a theory of spectral approximation for compact operators in Banach spaces. Their theory can be applied to the analysis of many numerical procedures for the spectral approximation of differential operators, T, such that $T + \lambda I$ has a compact inverse for some $\lambda \in \mathbb{C}$. Most of the differential systems in the theory of elasticity are in this class.

However, there are many differential systems of interest in mathematical physics which do not have compact resolvents. These operators can have continuous spectrum, eigenvalues of infinite multiplicity, and finite limit points of eigenvalues. Also, the eigenfunctions need not be smooth since the differential systems are not necessarily elliptic.

Descloux, Nassif, and Rappaz [4,5] have studied the approximation of the spectrum of a differential system of interest in magnetohydrodynamics which has a bounded inverse, but not a compact inverse. They developed a theory of spectral approximation for bounded operators which treats this problem. An analysis of the approximation of the spectral properties of a class of bounded operators by finite element methods has also been done by Mills [12, 13].

The results in this paper apply to closed (not necessarily bounded) operators in Banach spaces. We apply the perturbation theory developed in this paper to the study of a finite element procedure for approximating the spectral properties of a differential system modeling a fluid in a rotating basin. We note that unlike previous authors, we analyze the approximation of the differential operator directly and not through its inverse.

The time dependent equations for the differential system modeling a fluid in a rotating basin are

$$
(0.1)\qquad
\begin{aligned}
&\frac{\partial\zeta}{\partial t} = -\nabla \cdot \vec{u}\,, && (x,t) \in \bar{\Omega} \times \mathbb{R}\,,\\
&\frac{\partial\vec{u}}{\partial t} = -\nabla\zeta - f\vec{u} - \omega R\vec{u}\,, && (x,t) \in \bar{\Omega} \times \mathbb{R}\,,\\
&\vec{u} \cdot \vec{n} = 0\,, && (x,t) \in \partial\Omega \times \mathbb{R}\,,\\
&\int_\Omega \zeta\,dx = 0\,,
\end{aligned}
$$

where $\Omega \subseteq \mathbb{R}^2$ is a bounded, connected open set with smooth boundary, $\partial\Omega$, $\vec{u} = (u_1,u_2)$ is the horizontal volume transport, ζ is height of the fluid above equilibrium level, R is the linear operator $R\vec{u} = (-u_2,u_1)$, $\vec{n}$ is the exterior normal to $\partial\Omega$, and $f \geq 0$ and ω are real constants representing friction and Coriolis terms.

Thus, we are concerned with approximating the spectral properties of the system

$$
(0.2)\qquad
\begin{aligned}
&T(\zeta,\vec{u}) = (-\nabla \cdot \vec{u}, -\nabla\zeta - f\vec{u} - \omega R\vec{u})\,, && x \in \Omega\,,\\
&\vec{u} \cdot \vec{n} = 0\,, && x \in \partial\Omega\,,\\
&\int_\Omega \zeta\,dx = 0\,.
\end{aligned}
$$

We note that $T - I$ is a symmetric, formally dissipative operator with maximal non-positive boundary conditions [9] .

If $f = \omega = 0$, then $T(0, R\nabla\psi) = 0$ for all smooth functions ψ such that $\psi = 0$ on $\partial\Omega$. Hence, in this case $(f = \omega = 0)$, 0 is an eigenvalue of T of infinite multiplicity. Since the sum of an operator with non-compact inverse and a bounded operator has a non-compact inverse, it follows that $T - \lambda I$ has a non-compact inverse for the general case $f, \omega \in \mathbb{R}$, $\lambda \in \mathbb{C}$, λ belonging to the resolvent set of T .

We give here error estimates for a finite element procedure proposed by Platzman [15] to approximate the spectral properties of (0.2) . The selfadjoint case $f = 0$ was analyzed by Luskin [11] by techniques different from those used here.

In Section 1 , we give a general theory for the approximation of a closed operator A by a family of finite dimensional operators $\{A_h\}$. In our applications, A will be a differential operator and A_h will be an approximation of A given by a finite element procedure. We propose two properties P1) and P2) and show that these properties imply the convergence of the spectral properties of A_h to those of A . Error estimates in the applications will follow from Theorem 1.3.

We give special results in Section 2 for the case A and A_h are selfadjoint operators in a Hilbert space. These results apply to the approximation of continuous spectrum.

The operator theory developed in Sections 1 and 2 is applied in Section 3 and 4 to examples of the approximation of

spectral properties of differential operators by finite element methods. In Section 3, we apply our theory to obtain error estimates for the approximation of the spectral properties of scalar, second order, uniformly strongly elliptic operators by the standard finite element method. This example is included even though the results are not new since we believe that its inclusion will make it easier for the reader to understand our main example in Section 4.

In Section 4, we define and analyze the approximation of T. Theorem 1.3 gives optimal order estimates for the approximation of eigenspaces. Optimal order eigenvalue estimates for this problem have been given in [11] for the self-adjoint case and are derived from the results in this paper and [3] for the general case.

We note that we use without explicit reference the classical spectral theory; see for example Kato [6], Riesz-Nagy [16].

1. Approximation of the Spectrum of Closed Operators in Banach Spaces.

We first introduce some notation. Let X be a complex Banach space with norm $\| \ \|$. Denote by $\mathcal{B}(X)$ the set of bounded, linear operators $B: X \to X$. Also, denote by $\mathcal{C}(X)$ the set of closed, linear operators $C: \mathcal{D}(C) \subset X \to X$ where the domain of C, $\mathcal{D}(C)$, is not necessarily dense in X. For $C \in \mathcal{C}(X)$, $\rho(C)$ is the resolvent set of C defined by

$$\rho(C) = \{z \in \mathbb{C} \mid (z-C)^{-1} \in \mathcal{B}(X)\} .$$

If $z \in \rho(C)$, we define the resolvent operator $R_z(C) = (z-C)^{-1}$: $X \to X$. The complement of $\rho(C)$ is $\sigma(C) = \{z \in \mathbb{C} | z \notin \rho(C)\}$, the spectrum of C.

Let Y and Z be closed subspaces of X and $x \in X$. We set

$$\delta(x,Y) = \inf_{y \in Y} \|x-y\| , \quad \delta(Y,Z) = \sup_{\substack{y \in Y \\ \|y\|=1}} \delta(y,Z) ,$$

$$\hat{\delta}(Y,Z) = \max\big(\delta(Y,Z), \delta(Z,Y)\big) ;$$

$\hat{\delta}(Y,Z)$ is called the gap between Y and Z and is a measure of the "distance" between these spaces. If C and D are in $\mathcal{C}(X)$ with graphs $G_C, G_D \subset X \times X$, then we define $\delta(C,D) = \delta(G_C, G_D)$, i.e.,

$$\delta(C,D) = \sup_{\substack{x \in \mathcal{D}(C) \\ \|x\| + \|Cx\| = 1}} \ \inf_{y \in \mathcal{D}(D)} \{ \|x-y\| + \|Cx-Cy\| \} .$$

Furthermore, we define

$$\hat{\delta}(C,D) = \max\big(\delta(C,D), \delta(D,C)\big) .$$

Finally, if Y is subspace of $\mathcal{D}(C) \cap \mathcal{D}(D)$, we set

$$\|C-D\|_Y = \sup_{\substack{y \in Y \\ \|y\|=1}} \|Cy-Dy\| .$$

Now, let $A \in \mathcal{C}(X)$ be a given operator. In order to approximate $\sigma(A)$, we consider a family $\{X_h\}$ of finite dimensional subspaces of X parametrized by h and linear operators A_h: $X_h \to X_h$.

We denote by $\sigma(A_h)$, $\rho(A_h)$, and $R_z(A_h)$: $X_h \to X_h$ the spectrum, resolvent set, and resolvent operator of A_h considered as a bounded operator in X_h . However, when used in connection with expressions of the type "$\delta(A_h, A)$" or "$\|A-A_h\|_{X_h}$", A_h is considered as a closed operator in X with non dense domain X_h .

Let $\Gamma \subset \rho(A)$ be a given Jordan closed curve; then

$$E = \frac{1}{2\pi i} \int_\Gamma R_z(A)dz: \quad X \to X$$

and if $\Gamma \subset \rho(A_h)$,

$$E_h = \frac{1}{2\pi i} \int_\Gamma R_z(A_h)dz: \quad X_h \to X_h$$

are the spectral projectors relative to A and A_h respectively. We shall also use the relations

$$AE = \frac{1}{2\pi i} \int_\Gamma zR_z(A)dz: \quad X \to X$$

and

$$A_hE_h = \frac{1}{2\pi i} \int_\Gamma zR_z(A_h)dz: \quad X_h \to X_h .$$

We now introduce the following two properties:

P1) $\lim_{h\to 0} \delta(A_h, A) \to 0$,

P2) $\forall x \in X$, $\lim_{h\to 0} \delta(x, X_h) = 0$.

The following theorems contain the main results of this section. We note that $c < \infty$ shall denote in this paper a positive constant which is independent of h , but which varies from estimate to estimate.

<u>Theorem 1.1</u>. Suppose P1) is valid and let $K \subset \rho(A)$ be a compact set. Then there exists $h_0 > 0$ and $c = c(K)$ such that for $h < h_0$ we have $K \subset \rho(A_h)$ and

$$\|R_z(A) - R_z(A_h)\|_{X_h} \leq c\delta(A_h, A) , \qquad z \in K .$$

Theorem 1.1 shows that if P1) is valid, then the approximation of $\sigma(A)$ by $\sigma(A_h)$ is upper semicontinuous. Furthermore, for $h < h_0$, $\|R_z(A_h)\|_{X_h}$ is uniformly bounded on K . This is a stability property.

<u>Theorem 1.2</u>. Suppose P1) is valid. Then there exists $h_0 > 0$ and c so that for $h < h_0$ we have the bounds

(1.1) $$\|E - E_h\|_{X_h} + \|AE - A_h E_h\|_{X_h} \leq c\delta(A_h, A) ,$$

(1.2) $$\delta\big(E_h(X_h), E(X)\big) \leq c\delta(A_h, A) ,$$

(1.3) $$\delta\big(x, E_h(X_h)\big) \leq c\{\delta(x, X_h) + \delta(A_h, A)\,\|x\|\} , \quad x \in E(x) .$$

<u>Theorem 1.3</u>. Suppose that P1) and P2) are satisfied and that $E(X)$ is the finite dimensional subspace corresponding to an isolated eigenvalue λ of algebraic multiplicity m of A. Let α be the ascent of $(\lambda - A)$ and let f be a holomorphic function defined in the neighborhood of λ. We set

$$\gamma_h = \min\{\delta(A|_{E(X)}, A_h), \delta(A_h, A)\}$$

Then for h small enough, $A_h|_{E_h(X_h)}: E_h(X_h) \to E_h(X_h)$ has exactly m eigenvalues $\lambda_{1,h}, \ldots, \lambda_{m,h}$ repeated according to multiplicity. Also, there exists $h_0 > 0$ and c such that the following bounds are valid for $h < h_0$:

(1.4) $$\hat{\delta}\big(E_h(X_h), E(X)\big) \leq c\gamma_h ,$$

(1.5) $$\hat{\delta}\big(A_h|_{E_h(X_h)}, A|_{E(X)}\big) \leq c\gamma_h ,$$

(1.6) $$|f(\lambda) - \frac{1}{m} \sum_{i=1}^{m} f(\lambda_{i,h})| \leq c\gamma_h ,$$

(1.7) $$\max_{i=1,\ldots,m} |\lambda - \lambda_{i,h}|^{\alpha} \leq c\gamma_h .$$

<u>Remarks</u>. If P1) and P2) are satisfied, then (1.3) of Theorem 1.2 shows that $\lim_{h\to 0} \delta\big(x, E_h(X_h)\big) = 0$, $\forall x \in E(X)$, i.e., any $x \in E(X)$ can be approximated by vectors in $E_h(X_h)$. It is not true however that any $z \in \sigma(A)$ can be approximated by eigenvalues of A_h, i.e., there may exist $z \in \sigma(A)$ so that $\mathrm{dist}\big(z, \sigma(A_h)\big)$ fails to converge to zero. (the approximation of $\sigma(A)$ by $\sigma(A_h)$ is not necessarily lower semicontinuous). The classical counter-example is the shift operator (see, for

example, Kato [6, p. 210]). Note that if A and A_h are selfadjoint operators in a Hilbert space, then lower semicontinuity is true under weaker conditions than P1), P2) (see, for example, Kato [6, p. 431]).

In order to prove Theorems 1.1 and 1.2, we first prove a sequence of lemmas which are variants of results found in Kato [6, p. 197-208].

Lemma 1.1. Let $B \in \mathcal{B}(X)$, $C \in \mathcal{C}(X)$. Then

a) $$\delta(C,B) \le \|C-B\|_{\mathcal{D}(C)},$$

b) $$\|C-B\|_{\mathcal{D}(C)} \le (1+\|B\|)^2 \delta(C,B)/(1-(1+\|B\|)\delta(C,B))$$

if the denominator is positive.

Proof. Part a) follows directly from the definitions. We prove part b). Let $x \in \mathcal{D}(C)$, $\|x\| = 1$ and let $\varepsilon > 0$ be arbitrary. It follows from the definition of $\delta(C,B)$ that there exists $y \in X$ such that

$$\|x-y\| + \|Cx-By\| \le \delta(C,B)\{\|x\| + \|Cx\|\} + \varepsilon. \tag{1.8}$$

Consequently, we have

$$\begin{aligned} \|(C-B)x\| &\le \|Cx-By\| + \|By-Bx\| \\ &\le \|Cx - By\| + \|B\|\,\|y-x\| \\ &\le (1+\|B\|)\{\|Cx-By\| + \|y-x\|\} \\ &\le (1+\|B\|)\delta(C,B)\{\|x\| + \|Cx\|\} + \varepsilon(1+\|B\|). \end{aligned} \tag{1.9}$$

Replacing $\|Cx\|$ by $\|(C-B)x\| + \|B\| \|x\|$ in the right hand side of the above inequality and letting $\varepsilon \to 0$ yields the estimate

$$\|(C-B)x\| \le \tag{1.10}$$
$$(1+\|B\|)\delta(C,B)\{(1+\|B\|)\|x\| + \|(C-B)x\|\} .$$

The result follows directly from (1.10) Q.E.D.

Remark. It follows from Lemma 1.1 that if $A \in \mathcal{B}(x)$, then property P1) is equivalent to $\lim_{h\to 0}\|A-A_h\|_{X_h} = 0$. This is the spectral approximation condition of Descloux-Nassif-Rappaz [4].

Lemma 1.2. Let $B \in \mathcal{B}(X)$ and $C,D \in \mathcal{C}(X)$. Then

$$\delta(C+B,D+B) \le (1+\|B\|)^2\delta(C,D) .$$

Proof. Let $x \in \mathcal{D}(C)$, $\|x\| + \|(C+B)x\| = 1$ and let $\varepsilon > 0$ be arbitrary. We have by the triangle inequality that

$$\|x\| + \|Cx\| \le \|x\| + \|(C+B)x\| + \|Bx\| \le 1 + \|B\| . \tag{1.11}$$

It follows from the definition of $\delta(C,D)$ that we may choose $y \in \mathcal{D}(D)$ such that

$$\|x-y\| + \|Cx-Dy\| \le \delta(C,D)(1+\|B\|) + \varepsilon . \tag{1.12}$$

Hence,

$$\begin{aligned} \|x-y\| &+ \|(C+B)x - (D+B)y\| \\ &\le \|x-y\| + \|Cx-Dy\| + \|B\| \|x-y\| \\ &\le (1+\|B\|)^2\delta(C,D) + (1+\|B\|)\varepsilon . \end{aligned} \tag{1.13}$$

The result follows directly from (1.13) after letting $\varepsilon \to 0$.

Q.E.D.

Lemma 1.3. If C and $D \in \mathcal{C}(X)$ are invertible, then

$$\delta(C,D) = \delta(C^{-1},D^{-1}) .$$

Proof. This result follows directly from the definitions. Q.E.D.

Lemma 1.4. Let $C,D \in \mathcal{C}(X)$ and suppose there exists κ such that

$$\|Dx\| \geq \kappa \|x\|, \qquad x \in \mathcal{D}(D) . \tag{1.14}$$

If $\delta(C,D) < \min(1,\kappa)$, then C is invertible.

Proof. We show that if C is not invertible, then $\delta(C,D) \geq \min(1,\kappa)$. Let $x \in \mathcal{D}(C)$ be such that $\|x\| = 1$ and $Cx = 0$. Let $\varepsilon > 0$ be arbitrary. It follows from the definition of $\delta(C,D)$ that since $Cx = 0$ we can choose $y \in \mathcal{D}(D)$ such that

$$\|x-y\| + \|Dy\| \leq \delta(C,D) + \varepsilon .$$

It follows from (1.14) that

$$\begin{aligned} \delta(C,D) + \varepsilon &\geq \|x-y\| + \kappa \|y\| \\ &\geq |1- \|y\| | + \kappa \|y\| \\ &\geq \min(1,\kappa)\{|1- \|y\| | + \|y\| \} \\ &\geq \min(1,\kappa) . \end{aligned} \tag{1.15}$$

Since $\varepsilon > 0$ was arbitrary, it follows that we have reached the contradiction $\delta(C,D) \geq \min(1,\kappa)$. Q.E.D.

<u>Proof of Theorem 1.1.</u> In this proof, c depends on K but all estimates are uniform for $z \in K$. By Lemma 1.2,

$$\delta(z-A_h, z-A) \leq c\delta(A_h, A), \qquad z \in K. \tag{1.16}$$

It follows from (1.16) and P1) that

$$\lim_{h \to 0} \delta(z-A_h, z-A) \to 0 \text{ uniformly for } z \in K. \tag{1.17}$$

Since $K \subset \rho(A)$, there exists c_1 such that

$$\|(z-A)x\| \geq c_1 \|x\|, \qquad x \in \mathcal{D}(A), \quad z \in K. \tag{1.18}$$

We can now conclude from (1.17), (1.18) and Lemma 1.4 that there exists $h_0 > 0$ such that $z - A_h$ is invertible for $h < h_0$ and $z \in K$. Since X_h is finite dimensional, we have that $K \subset \rho(A_h)$ for $h < h_0$.

Furthermore, it follows from (1.16) and Lemma 1.3 that

$$\delta\big(R_z(A_h), R_z(A)\big) \leq c\delta(A_h, A), \qquad z \in K. \tag{1.19}$$

Hence, we can obtain from Lemma 1.1b the result

$$\begin{aligned} \|R_z(A)-R_z(A_h)\|_{X_h} &\leq c\delta\big(R_z(A_h), R_z(A)\big) \\ &\leq c\delta(A_h, A), \qquad h < h_0,\ z \in K. \end{aligned} \tag{1.20}$$

Q.E.D.

<u>Proof of Theorem 1.2.</u> The result (1.1) follows from Theorem 1.1 and the following estimates:

$$
\begin{aligned}
&\|E-E_h\|_{X_h} + \|AE-A_hE_h\|_{X_h} \\
(1.21)\qquad &= \left\|\frac{1}{2\pi i}\int_\Gamma \big(R_z(A)-R_z(A_h)\big)dz\right\|_{X_h} \\
&+ \left\|\frac{1}{2\pi i}\int_\Gamma \big(zR_z(A) - z\,R_z(A_h)\big)dz\right\|_{X_h} \\
&\leq (2\pi)^{-1} \max_{z\in\Gamma} \|R_z(A)-R_z(A_h)\|_{X_h} \int_\Gamma (1+|z|)|dz| .
\end{aligned}
$$

The estimate (1.2) now follows directly from the estimate (1.1).

In order to prove (1.3), let $x \in E(X)$ and $x_h \in X_h$. Then $x - E_hx_h = E(x-x_h) + (E-E_h)x_h$. Consequently,

$$
\begin{aligned}
&\delta\big(x,E_h(X_h)\big) \leq \big(\|E\|\ \|x-x_h\| + \|E-E_h\|_{X_h} \|x_h\| \big) \\
(1.22)\qquad &\leq \big(\|E\| + \|E-E_h\|_{X_h}\big) \|x-x_h\| + \|E-E_h\|_{X_h} \|x\| .
\end{aligned}
$$

The result (1.3) follows by taking the infimum over $x_h \in X_h$ and using (1.1). Q.E.D.

It remains to prove Theorem 1.3. We first quote without proof the following simple result:

<u>Lemma 1.5.</u> Let Y and Z be two subspaces of X with the same finite dimension and let $P\colon Y \to Z$ be a linear operator such that

$$(1.23)\qquad \|Py-y\| \leq \tfrac{1}{2}\|y\| , \qquad y \in Y .$$

Then P is bijective and

$$(1.24)\qquad \|P^{-1}z\| \leq 2\,\|z\| , \qquad z \in Z .$$

Lemma 1.6. Let Y and Z be two subspaces of X .

a) If $\hat{\delta}(Y,Z) < 1$, then $\dim Y = \dim Z$.

b) If $\dim Y = \dim Z < \infty$

$$\delta(Y,Z) \leq \delta(Z,Y)[1-\delta(Z,Y)]^{-1}$$

Proof. For a) , See Kato [6, p. 200] . For b , see Kato [7]. Q.E.D.

In this rest of this section, we suppose that the hypotheses of Theorem 1.3 are satisfied. Also, c and h_0 will denote two generic positive constants which depend on Γ .

Lemma 1.7. $\delta\left(A|_{E(x)}, A_h|_{E_h(x_h)}\right) \leq c\delta(A|_{E(x)}, A_h)$, $h \leq h_0$.

Proof. Let $x \in E(x)$, $\|x\| + \|Ax\| = 1$, and let $x_h \in X_h$ be such that

$$\|x-x_h\| + \|Ax-A_hx_h\| \leq \delta\left(A|_{E(x)}, A_h\right) .$$

We have

$$\begin{aligned} &\|x-E_hx_h\| + \|Ax-A_hE_hx_h\| \\ &= \left\|\frac{1}{2\pi i}\int_\Gamma \left(R_z(A)x-R_z(A_h)x_h\right)dz\right\| \\ (1.25)\qquad &+ \left\|\frac{1}{2\pi i}\int_\Gamma \left(zR_z(A)x - zR_z(A_h)x_h\right)dz\right\| \\ &\leq (2\pi)^{-1}\max_{z\in\Gamma}\|R_z(A)x-R_z(A_h)x_h\| \int_\Gamma (1+|z|)|dz| . \end{aligned}$$

In order to conclude the proof of the lemma, it suffices to estimate $\|R_z(A)x-R_z(A_h)x_h\|$ for $z \in \Gamma$. We set $w = R_z(A)x \in E(X)$ and let $w_h \in X_h$ be such that

$$(1.26) \qquad \|w-w_h\| + \|Aw-A_hw_h\| \le \delta(A|_{E(x)},A_h)(\|w\| + \|Aw\|).$$

However,

$$\begin{aligned} \|w\| + \|Aw\| &= \|R_z(A)x\| + \|R_z(A)Ax\| \\ &\le \|R_z(A)\| \left(\|x\| + \|Ax\|\right) \\ &= \|R_z(A)\| \le c, \qquad z \in \Gamma . \end{aligned}$$

Thus,

$$(1.27) \qquad \|w-w_h\| + \|Aw-A_hw_h\| \le c\delta(A|_{E(x)},A_h) .$$

By using Theorem 1.1, i.e., $\|R_z(A_h)\|_{X_h} \le c, \quad h \le h_0, \quad z \in \Gamma$, we have

$$(1.28) \qquad \begin{aligned} \|R_z(A_h)x_h-w_h\| &= \|R_z(A_h)\left(x_h-(z-A_h)w_h\right)\| \\ &\le c\{\|x_h-x\| + \|(z-A)w-(z-A_h)w_h\|\} \\ &\le c\delta(A|_{E(x)},A_h) . \end{aligned}$$

Consequently,

$$(1.29) \qquad \begin{aligned} \|R_z(A)x - R_z(A_h)x_h\| &\le \|w-w_h\| + \|w_h-R_z(A_h)x_h\| \\ &\le c\delta(A|_{E(x)},A_h) . \end{aligned}$$

Q.E.D.

<u>Remark</u>. Lemma 1.7 is still valid if the hypothesis P1) is replaced by the uniform boundedness of $R_z(A_h)$ on Γ.

Proof of Theorem 1.3. From (1.2) and (1.3) of Theorem 1.2 it follows that $\lim_{h\to 0} \hat{\delta}(E_h(X_h),E(X)) = 0$. Consequently, by Lemma 1.6 a), $\dim E_h(X_h) = \dim E(X) = m$ for $h < h_0$. This proves that for $h < h_0$, $A_h|_{E_h(X_h)}: E_h(X_h) \to E_h(X_h)$ has exactly m eigenvalues $\lambda_{1,h},\ldots,\lambda_{m,h}$ repeated according to multiplicity.

By Lemma 1.6 b), we have that

$$\begin{aligned} \hat{\delta}(E(X),E_h(X_h)) &\le c \min\{\delta(E(X),E_h(X_h)) , \\ &\qquad \delta(E_h(X_h),E(X))\} , \qquad h < h_0 . \end{aligned} \tag{1.30}$$

It follows from Lemma 1.7 that

$$\begin{aligned} \delta(E(X), E_h(X_h)) &\le c\delta(A|_{E(X)},A_h|_{E_h(X_h)}) \\ &\le c\delta(A|_{E(X)},A_h) . \end{aligned} \tag{1.31}$$

The result (1.31) and (1.2) of Theorem 1.2 yield (1.4) when substituted in (1.30) .

We have by (1.1) of Theorem 1.2 that

$$\delta(A_h|_{E_h(X_h)} , A|_{E(X)}) \le c\delta(A_h,A), \qquad h < h_0 . \tag{1.32}$$

Hypotheses P1) and Lemmas 1.6 and 1.7 allow us to conclude the validity of (1.5) .

Now let $u_1,\ldots,u_m$ be a basis for $E(X)$ with $\|u_i\| + \|Au_i\| = 1$, $i = 1,\ldots,m$. We then choose $u_{1,h},\ldots,u_{m,h} \in E_h(X_h)$ so that

$$\|u_i - u_{i,h}\| + \|Au_i - A_h u_{i,h}\| \le \delta(A|_{E(X)}, A_h|_{E_h(X_h)}) , \tag{1.33}$$
$$i = 1,\dots,m .$$

Next, we define $\Lambda_h\colon\ E(X) \to E_h(X_h)$ as the linear operator such that $\Lambda_h u_i = u_{i,h}$, $i = 1,\dots,m$. It follows by (1.5) that

$$\|u - \Lambda_h u\| + \|Au - A_h \Lambda_h u\| \le c\gamma_h \|u\| , \qquad u \in E(X) . \tag{1.34}$$

By P1) $\lim\limits_{h\to 0} \gamma_h = 0$, so by Lemma 1.5 and (1.34) Λ_h is a bijection whose inverse Λ_h^{-1} is uniformly bounded for $h < h_0$. Let $\hat{A} = A|_{E(X)}$ and $\hat{A}_h = \Lambda_h^{-1} A_h \Lambda_h\colon\ E(X) \to E(X)$. For $u \in E(X)$, we have

$$\begin{aligned} \|(\hat{A} - \hat{A}_h)u\| &= \|\Lambda_h^{-1}(\Lambda_h A - A_h \Lambda_h)u\| \\ &\le c\{ \|(\Lambda_h - I)Au\| + \|Au - A_h \Lambda_h u\| \} \\ &\le c\gamma_h \|u\| , \qquad h < h_0 . \end{aligned} \tag{1.35}$$

Consequently,

$$\|\hat{A} - \hat{A}_h\|_{E(X)} \le c\gamma_h . \tag{1.36}$$

Now $\hat{A}_h$ has eigenvalues $\lambda_{1,h},\dots,\lambda_{m,h}$ and $\hat{A}$ has the eigenvalue λ of algebraic multiplicity m . Also, α is the ascent of $(\lambda - \hat{A})$. We have reduced our problem to the matrix case. The results (1.6) and (1.7) of Theorem 1.3 now follow from the classical perturbation theory for finite dimensional operators [17, p. 80-81]. Q.E.D.

2. The Approximation of the Spectrum of a Selfadjoint Operator.

In this section, we suppose that X is a Hilbert space, $A: \mathscr{D}(A) \subset X \to X$ is selfadjoint, and $A_h: X_h \to X_h$ is selfadjoint in X_h for all h. If $I \subset \mathbb{R}$ is an interval (finite or infinite), $E_I: X \to X$ will denote the spectral projector of A relative to I and $E_{h,I}: X_h \to X_h$ will denote the spectral projector of A_h relative to I. We can prove the following theorems:

<u>Theorem 2.1.</u> P1) $\Longleftrightarrow$ $\forall$ closed intervals I,J, one of them bounded, $I \cap J = \phi$, we have

$$\lim_{h\to 0} \|E_I E_{h,J}\|_{X_h} = 0 .$$

<u>Theorem 2.2.</u> Suppose that P1) is valid. Let $J \subset I$ where J is a closed bounded interval and I is an open interval. Then

$$\lim_{h\to 0} \delta\big(E_{h,J}(X_h), E_I(X)\big) = 0 . \tag{2.1}$$

<u>Theorem 2.3.</u> Suppose that P1) and P2) are valid.

a) Let I be an open interval and $x \in E_I(X)$. Then

$$\lim_{h\to 0} \delta\big((x, E_{h,I}(X_h)\big) = 0 .$$

b) If $\lambda \in \sigma(A)$, then

$$\lim_{h\to 0} \operatorname{dist}\big(\lambda, \sigma(A_h)\big) = 0$$

Remark. Theorem 2.3b states that the approximation of $\sigma(A)$ by $\sigma(A_h)$ is lower semicontinuous.

Theorems 2.1, 2.2, and 2.3 have been proved in Section 3 of Descloux, Nassif, Rappaz [4] when A is bounded (recall our remarks after Lemma 1.1). Our proof will reduce the unbounded case to the bounded case. We shall restrict ourself to the proof of Theorem 2.2. Theorems 2.1 and 2.3 can be obtained by similar arguments. Furthermore, for the sake of simplicity and without much loss of generality we shall suppose that $\rho(A) \cap \mathbb{R} \neq \phi$.

Proof of Theorem 2.2. As mentioned above, we suppose there exists $a \in \mathbb{R}$ with $a \in \rho(A)$. We introduce the function $\phi(\lambda) = (a-\lambda)^{-1}$ and set $B = \phi(A) = R_a(A)$. By Theorem 1.1, $B_h = \phi(A_h) = R_a(A_h)$ is well-defined for h sufficiently small and

$$\lim_{h\to 0} \|B-B_h\|_{X_h} = 0 .$$

For an interval M , $F_M\colon X \to X$ and $F_{h,M}\colon X_h \to X_h$ will denote respectively the spectral projectors of B and B_h relative to M . As a first case, suppose $a \notin I$ and set $K = \phi(J)$ and $L = \phi(I)$. Then K and L are respectively a compact and an open interval with $K \subset L$. From Theorem 4 of Descloux, Nassif, Rappaz [4], we can conclude that

$$\lim_{h\to 0} \delta\big(F_{h,K}(X_h), F_L(X)\big) = 0 . \tag{2.2}$$

However, the result (2.1) follows from (2.2) since $F_{h,K}(X_h) = E_{h,J}(X_h)$ and $F_L(X) = E_I(X)$.

The case $a \in I$ can be reduced to the preceeding one by noting that we can find compact intervals J1, J2 and open intervals I1, I2 with the following properties (for $\hbar$ sufficiently small):

$$a \notin I1 \cup I2 ,$$

$$J1 \subset I1 , \quad J2 \subset I2 ,$$

$$E_I(X) = E_{I1}(X) \oplus E_{I2}(X) ,$$

$$E_{h,J}(X_h) = E_{h,J1}(X_h) \oplus E_{h,J2}(X_h) .$$

Q.E.D.

3. Application to Scalar Elliptic Boundary Value Problems.

We shall need the following notation to discuss the application of our theory to the approximation of the spectral properties of scalar elliptic boundary value problems. Denote by Ω an open set in $\mathbb{R}^n$ with a smooth boundary, $\partial\Omega$. As usual, we denote by $L^2(\Omega)$ the Hilbert space of square integrable, complex-valued functions with inner product and norm

$$\langle F,G\rangle = \int_\Omega F\bar{G}dx , \quad \|F\|^2 = \langle F,F\rangle . \tag{3.1}$$

We denote by $H^r(\Omega)$ the space of complex-valued functions whose distribution derivatives of order less than or equal to r, r a nonnegative integer, are in $L^2(\Omega)$ with norm

$$\|F\|_r^2 = \sum_{|\alpha| \le r} \|D^\alpha F\|^2 . \tag{3.2}$$

We wish to consider the spectral approximation of the operator

$$\begin{aligned} Lu &= -\sum_{i,j=1}^n D_i(a_{ij}D_ju) + \sum_{i=1}^n a_iD_iu + au , & x \in \Omega , \\ \mathcal{B}u &= \sum_{i,j=1}^n a_{ij}\nu_iD_ju = 0 , & x \in \partial\Omega , \end{aligned} \tag{3.3}$$

where we assume that L is a uniformly, strongly elliptic operator with real-valued coefficients in $C^\infty(\bar{\Omega})$ and $\nu = (\nu_1,\ldots,\nu_n)$ is the unit exterior normal to $\partial\Omega$. We associate with L the continuous, sesquilinear form on $H^1(\Omega) \times H^1(\Omega)$,

$$B(\phi,\psi) = \sum_{i,j=1}^{n} \langle a_{ij} D_j \phi, D_i \psi \rangle + \sum_{i=1}^{n} \langle a_i D_i \phi, \psi \rangle \tag{3.4}$$
$$+ \langle a\phi, \psi \rangle$$

We may assume (by replacing L by $L + \mu$, $\mu \in \mathbb{R}$) that for some $b > 0$

$$\text{Re}\, B(\phi,\phi) \geq b \|\phi\|_1^2 , \qquad \phi \in H^1(\Omega) . \tag{3.5}$$

Let $\mathcal{S}_h \subset H^1(\Omega)$ be a family of finite dimensional subspaces parametrized by h , $0 < h \leq 1$, with r a positive integer and c a positive constant, independent of h , such that for $1 \leq s \leq r + 1$, $u \in H^s(\Omega)$, we have

$$\inf_{\chi \in \mathcal{S}_h} (\|u-\chi\| + h \|u-\chi\|_1) \leq ch^s \|u\|_s . \tag{3.6}$$

Many finite element spaces are known to satisfy (3.6) [2] .

We define the operator L_h: $\mathcal{S}_h \to \mathcal{S}_h$ by the relation

$$B(U_h,W) = \langle L_h U_h, W \rangle , \qquad W \in \mathcal{S}_h . \tag{3.7}$$

Note that L_h is well-defined since $\mathcal{S}_h$ is finite dimensional. We shall study the approximation of the spectral properties of L by the spectral properties of L_h .

Define the space

$$H^2_{\mathcal{B}}(\Omega) = \{u \in H^2(\Omega) \,|\, \mathcal{B} u = 0 \text{ for } x \in \partial\Omega\} .$$

We consider L as a closed operator from $L^2(\Omega)$ to $L^2(\Omega)$, i.e., $X = L^2(\Omega)$, with domain $D(L) = H^2_{\mathcal{B}}(\Omega)$. Also, in the

notation of Section 1, we set $X_h = \mathcal{S}_h$. It follows from (3.6) that if $u \in C^\infty(\bar{\Omega})$, then

$$\delta(u, \mathcal{S}_h) = \inf_{\chi \in \mathcal{S}_h} \|u-\chi\| \to 0 \quad \text{as} \quad h \to 0 .$$

Since $C^\infty(\bar{\Omega})$ is dense in $L^2(\Omega)$, it is clear that property P2) is valid.

We now turn to the verification of P1). To that end, let $P_h^o\colon\ L^2(\Omega) \to \mathcal{S}_h$ be the $L^2(\Omega)$ projection,

$$(3.8) \qquad \langle P_h^o u-u, W\rangle = 0 , \qquad W \in \mathcal{S}_h ,$$

and let $P_h^1\colon\ H^1(\Omega) \to \mathcal{S}_h$ be the $H^1(\Omega)$ projection defined by

$$(3.9) \qquad B(P_h^1 u-u, W) = 0 , \qquad W \in \mathcal{S}_h .$$

It follows from (3.5) that P_h^1 is well-defined.

The following estimates for P_h^o and P_h^1 are well-known [2]. There exists $c < \infty$ such that for $0 \le s \le r + 1$ and $u \in H^s(\Omega)$,

$$(3.10) \qquad \|P_h^o u-u\| \le ch^s \|u\|_s$$

and such that for $1 \le s \le r + 1$ and $u \in H^s(\Omega)$,

$$(3.11) \qquad \|P_h^1 u-u\| + h\|P_h^1 u-u\|_1 \le ch^s \|u\|_s .$$

We now verify that

$$(3.12) \qquad \delta(L_h, L) \to 0 .$$

For $U_h \in \mathcal{S}_h$, let $u \in D(L)$ satisfy

$$Lu = L_h U_h \quad . \tag{3.13}$$

It follows by elliptic regularity that

$$\|u\|_2 \le c \, \|L_h U_h\| \quad . \tag{3.14}$$

Now it is easily verified that $U_h = P_h^1 u$. Hence, by (3.11) and (3.14) ,

$$\|u - U_h\| = \|u - P_h^1 u\| \le ch^2 \|u\|_2 \le ch^2 \|L_h U_h\| \quad . \tag{3.15}$$

Since $Lu = L_h U_h$, we can conclude that for each $U_h \in \mathcal{S}_h$, there exists $u \in D(L)$ such that

$$\|u - U_h\| + \|Lu - L_h U_h\| \le ch^2 (\|U_h\| + \|L_h U_h\|) \quad . \tag{3.16}$$

Hence, we can conclude that

$$\delta(L_h, L) \le ch^2 \quad .$$

It is well-known that the spectrum of L consists of isolated eigenvalues with finite dimensional generalized eigenspaces. Let $\lambda \in \sigma(L)$ and let $E(X)$ be the generalized eigenspace corresponding to λ . Thus, $\dim E(X) < \infty$. We shall show that

$$\delta(L|_{E(X)}, L_h) \le ch^{r+1} \quad . \tag{3.17}$$

The result (3.17) will imply that the conclusions of Theorem 1.3 are valid with $\gamma_h \le ch^{r+1}$.

If $u \in E(X)$, set $U_h = P_h^1 u \in \mathcal{S}_h$. It follows from elliptic regularity that $E(X) \subset C^\infty(\bar{\Omega})$. Hence, by (3.11)

$$\|u - P_h^1 u\| \le ch^{r+1} \|u\|_{r+1} . \tag{3.18}$$

Also, note that if $u \in H^2(\Omega)$, then

$$P_h^0 Lu = L_h P_h^1 u . \tag{3.19}$$

Thus, it follows from (3.10) that

$$\begin{aligned} \|Lu - L_h U_h\| &= \|Lu - L_h P_h^1 u\| \\ &= \|Lu - P_h^0 Lu\| \le ch^{r+1} \|Lu\|_{r+1} \end{aligned} \tag{3.20}$$

Note that since $\dim E(X) < \infty$ and all norms on finite dimensional spaces are equivalent, there exists a constant $c < \infty$ such that for $u \in E(X)$,

$$\|Lu\|_{r+1} \le c \|u\|$$

$$\|u\|_{r+1} \le c \|u\| .$$

So, (3.17) follows from (3.18) and (3.20).

The results on eigenvalue approximation implied by (3.17) are not the optimal $O(h^{2r})$ results for this problem that have been obtained by other methods [1,14] unless $r = 1$. However, it can be shown that optimal $O(h^{2r})$ eigenvalue results for the problem can be derived from Theorem 1.3 for this problem if the approximation of L by L_h is analyzed in a space X which is taken to be an appropriate negative order Sobolev space.

4. Application to the Linearized Shallow Water Equations.

We turn now to the description and analysis of an approximation procedure for the spectral properties of the operator , T , associated with the linearized shallow water equations. Recall that

$$(4.1)\qquad \begin{aligned} &T(\zeta,\vec{u}) = (-\nabla\cdot\vec{u},-\nabla\zeta-f\vec{u}-\omega R\vec{u}) , && x \in \Omega , \\ &\vec{u}\cdot\vec{n} = 0 , && x \in \partial\Omega , \\ &\int_\Omega \zeta dx = 0 && \end{aligned}$$

where $\Omega \subseteq \mathbb{R}^2$ is a bounded, connected open set with smooth boundary, $\partial\Omega$, $\vec{u} = (u_1,u_2)$, R is the linear operator $R\vec{u} = (-u_2,u_1)$, $\vec{n}$ is the exterior normal to $\partial\Omega$, and $f \geq 0$ and ω are real constants representing friction and Coriolis terms. We assume that $\partial\Omega$ has a finite number of connected components, $\{\partial\Omega_i\}_{i=0}^s$. Also, assume that the sets $\partial\Omega_i$ are smooth arcs in $\mathbb{R}^2$.

We shall need the definitions of the following spaces of scalar-valued functions:

$$(4.2)\qquad \begin{aligned} L^2_*(\Omega) &= \{w \in L^2(\Omega) \mid \int_\Omega w dx = 0\} , \\ H^1_*(\Omega) &= H^1(\Omega) \cap L^2_*(\Omega) , \\ H^1_c(\Omega) &= \{w \in H^1(\Omega) \mid w(x) = 0 \text{ for } x \in \partial\Omega_0 \\ &\quad \text{and there exists constants, } (c_i)_{i=1}^s \text{, such that } w(x) = c_i \\ &\quad \text{for } x \in \partial\Omega_i,\ i = 1,\dots,s\} . \end{aligned}$$

For $k \geq 1$, set

$$(4.3)\qquad \begin{aligned} H^k_*(\Omega) &= H^1_*(\Omega) \cap H^k(\Omega) , \\ H^k_c(\Omega) &= H^1_c(\Omega) \cap H^k(\Omega) . \end{aligned}$$

We give these subspaces of $L^2(\Omega)$ and $H^1(\Omega)$ the inner product and norms defined in (3.1) and (3.2).

We also define the following spaces of vector-valued functions:

$$L^2(\Omega)^2 = \{\vec{u} = (u_1,u_2) \mid u_1,u_2 \in L^2(\Omega)\} ,$$
$$H^k(\Omega)^2 = \{\vec{u} = (u_1,u_2) \mid u_1,u_2 \in H^k(\Omega)\} \text{ for } k \geq 1.$$

With these spaces we associate the inner product and norms

$$\langle \vec{u},\vec{v} \rangle = \langle u_1,v_1 \rangle + \langle u_2,v_2 \rangle ,$$

$$\| \vec{u} \|^2 = \langle \vec{u},\vec{u} \rangle ,$$

$$\text{for } \vec{u} = (u_1,u_2), \quad \vec{v} = (v_1,v_2) \in L^2(\Omega)^2 ,$$

$$\| \vec{u} \|_k^2 = \| u_1 \|_k^2 + \| u_2 \|_k^2$$

$$\text{for } \vec{u} = (u_1,u_2) \in H^k(\Omega)^2 .$$

Finally, for functions $(\zeta,\vec{u})$, $(\xi,\vec{v}) \in L_*^2(\Omega) \times L^2(\Omega)^2$ we define the inner product

$$\langle (\zeta,\vec{u}), (\xi,\vec{v}) \rangle = \langle \zeta,\xi \rangle + \langle \vec{u},\vec{v} \rangle$$

and norm

$$\| (\zeta,u) \|^2 = \langle (\zeta,\vec{u}),(\zeta,\vec{u}) \rangle$$

and for functions $(\zeta,\vec{u}) \in H_*^1(\Omega) \times L^2(\Omega)^2$ we define the norm

$$\| (\zeta,u) \|_X^2 = \| \zeta \|_1^2 + \| \vec{u} \|^2 .$$

If $\vec{u} \in L^2(\Omega)^2$, then there exists unique scalar functions $\phi \in H^1_*(\Omega)$ and $\psi \in H^1_c(\Omega)$ such that

(4.4) $$\vec{u} = -\nabla\phi + R\nabla\psi \ .$$

The functions $\{\phi,\psi\}$ are known as Stokes-Helmholtz potentials for $\vec{u}$ [8,15], and we define the functions $S\vec{u} = \phi$, $H\vec{u} = \psi$. It follows from elliptic regularity [10] that for $k \geq 0$, there exists $c = c(k)$, $c < \infty$, such that

(4.5) $$\|S\vec{u}\|_{k+1} \leq c\|\vec{u}\|_k \ , \|H\vec{u}\|_{k+1} \leq c\,\|\vec{u}\|_k \ , \ \forall \vec{u} \in H^k(\Omega)^2 \ .$$

It is easily verified that

$$\langle \nabla\phi, R\nabla\psi \rangle = 0, \qquad \phi \in H^1_*(\Omega),\ \psi \in H^1_c(\Omega) \ . \tag{4.6}$$

Hence, (4.4) is an orthogonal direct sum decomposition.

Let $\{\mathcal{M}_h\}$, $0 < h \le 1$, be a family of finite dimensional subspaces of $H^1(\Omega)$ parametrized by h. Set

$$\mathcal{M}_h^c = \mathcal{M}_h \cap H^1_c(\Omega) \ , \tag{4.7}$$

$$\mathcal{M}_h^* = \mathcal{M}_h \cap H^1_*(\Omega) \ . \tag{4.8}$$

We assume that there exists a positive constant c, independent of h, and a positive integer r such that for $1 \le k \le r+1$ and $w \in H^k_*(\Omega)$,

$$\inf_{\chi \in \mathcal{M}_h^*} \left(\|w-\chi\| + h\|w-\chi\|_1 \right) \le ch^k \|w\|_k \ , \tag{4.9}$$

and for $1 \le k \le r+1$ and $w \in H^k_c(\Omega)$,

$$\inf_{\chi \in \mathcal{M}_h^c} \left(\|w-\chi\| + h\|w-\chi\|_1 \right) \le ch^k \|w\|_k \ . \tag{4.10}$$

We also wish to assume that the spaces $\{\mathcal{M}_h\}$ satisfy the "inverse property" that there exists a positive constant c, independent of h, such that

$$\|\chi\|_1 \le ch^{-1} \|\chi\|_0 \ , \quad \chi \in \mathcal{M}_h \ .$$

We shall use the spaces of vector-valued functions

$$\mathcal{N}_h = \{\vec{U} \mid \vec{U} = -\nabla\Phi + R\nabla\Psi,\ \Phi \in \mathcal{M}_h^*,\ \ \Psi \in \mathcal{M}_h^c\}$$

It follows from (4.5), (4.9), and (4.10) that for $0 \le k \le r$, $\vec{u} \in H^k(\Omega)^2$,

$$\inf_{\vec{\chi} \in \mathcal{N}_h} \|\vec{u}-\vec{\chi}\| \le \hat{c}h^k \|\vec{u}\|_k \quad (4.11)$$

We shall need the result that if $\Psi \in \mathcal{M}_h^c$, then

$$\nabla\Psi \in \mathcal{N}_h . \tag{4.12}$$

It is easy to see that (4.12) is valid if $\mathcal{M}_h$ contains the constant functions.

In order to define our approximation procedure, we introduce the continuous, sesquilinear form $B(\cdot,\cdot)$ on $X \times X = [H^1_*(\Omega) \times L^2(\Omega)^2] \times [H^1_*(\Omega) \times L^2(\Omega)^2]$ by

$$\begin{aligned} B\big((\zeta,\vec{u}),\ (\xi,\vec{v})\big) &= \langle \vec{u},\nabla\xi\rangle - \langle\nabla\zeta,\vec{v}\rangle \\ &\quad - \langle f\vec{u}+\omega R\vec{u},\vec{v}\rangle, \\ &\qquad (\zeta,\vec{u}),(\xi,\vec{v}) \in H^1_*(\Omega) \times L^2(\Omega)^2 . \end{aligned} \tag{4.13}$$

We regard T in (4.1) as an unbounded, closed operator, $T\colon\ H^1_*(\Omega) \times L^2(\Omega)^2 \to H^1_*(\Omega) \times L^2(\Omega)^2$, with domain

$$\begin{aligned} D(T) = \{ (\zeta,\vec{u}) \mid \zeta &\in H^1_*(\Omega) \\ \vec{u} &\in L^2(\Omega)^2, \nabla\cdot\vec{u} \in H^1(\Omega), \\ \vec{u}\cdot\vec{n} &= 0 \text{ on } \partial\Omega \} . \end{aligned} \tag{4.14}$$

Thus, if $(\zeta,\vec{u}) \in D(T)$,

$$\begin{aligned} B\big((\zeta,\vec{u}),(\xi,\vec{v})\big) &= \langle T(\zeta,\vec{u}),(\xi,\vec{v})\rangle, \\ &\qquad (\xi,\vec{v}) \in H^1_*(\Omega) \times L^2(\Omega)^2 . \end{aligned} \tag{4.15}$$

We now define $T_h\colon \mathcal{M}_h^* \times \mathcal{N}_h \to \mathcal{M}_h^* \times \mathcal{N}_h$ by

$$\begin{aligned} B\big((Z,\vec{U}),(Y,\vec{V})\big) &= \langle T_h(Z,\vec{U}),(Y,\vec{V})\rangle, \\ &\qquad (Y,\vec{V}) \in \mathcal{M}_h^* \times \mathcal{N}_h . \end{aligned} \tag{4.16}$$

The operator T_h is well-defined since the spaces $\mathcal{M}_h^* \times \mathcal{N}_h$ are finite dimensional.

We now show that the spectral properties of T_h approximate those of T by verifying P1) and P2) in this case for $X = H_*^1(\Omega) \times L^2(\Omega)^2$ and $X_h = \mathcal{M}_h^* \times \mathcal{N}_h$. The verification of P2) follows directly from (4.9) and (4.11). To see this, we note that for $(\zeta,\vec{u}) \in L_*^2(\Omega) \cap C^\infty(\bar{\Omega}) \times C^\infty(\bar{\Omega})^2$ it follows from (4.9) and (4.11) that

$$\delta\big((\zeta,\vec{u}),X_h\big) \to 0 \quad \text{as} \quad h \to 0 \ .$$

The validity of P2) now follows from the density of $L^2_*(\Omega) \cap C^\infty(\bar{\Omega}) \times C^\infty(\bar{\Omega})^2$ in $H^1_*(\Omega) \times L^2(\Omega)^2$.

We next turn to the verification of P1). In order to prove this result, we introduce and analyze some projection operators. First, we define the $H^1_*(\Omega)$ projection $P^*_h: H^1_*(\Omega) \to \mathcal{M}^*_h$ by

$$\langle \nabla(P^*_h w - w), \nabla\chi \rangle = 0, \qquad \chi \in \mathcal{M}^*_h, \tag{4.17}$$

and the $H^1_c(\Omega)$ projection $P^c_h: H^1_c(\Omega) \to \mathcal{M}^c_h$ by

$$\langle \nabla(P^c_h w - w), \nabla\chi \rangle = 0, \qquad \chi \in \mathcal{M}^c_h. \tag{4.18}$$

It is well-known [2] that there exists $c < \infty$ such that for $1 \le k \le r+1$ and $w \in H^k_*(\Omega)$,

$$\|P^*_h w - w\|_1 \le ch^{k-1}\|w\|_k \tag{4.19}$$

and for $1 \le k \le r+1$ and $w \in H^k_c(\Omega)$,

$$\|P^c_h w - w\|_1 \le ch^{k-1}\|w\|_k. \tag{4.20}$$

It follows from (4.11) that if $\vec{Q}_h: L^2(\Omega)^2 \to \mathcal{N}_h$ is the $L^2(\Omega)^2$ projection defined by

$$\langle \vec{Q}_h\vec{u} - \vec{u}, \vec{\chi} \rangle = 0, \qquad \vec{\chi} \in \mathcal{N}_h, \tag{4.21}$$

then there exists a constant $c < \infty$ such that $0 \le k \le r$ and $\vec{u} \in H^k(\Omega)^2$ implies that

$$\|\vec{Q}_h\vec{u} - \vec{u}\| \le ch^k\|\vec{u}\|_k. \tag{4.22}$$

It is easily checked that

$$\vec{Q}_h\vec{u} = -\nabla P_h^* S\vec{u} + R\nabla P_h^c H\vec{u} \quad . \tag{4.23}$$

Finally, define the $L^2(\Omega)$ projection $\hat{Q}_h$: $L_*^2(\Omega) \to \mathcal{M}_h^*$ by

$$\langle \hat{Q}_h w - w, \chi \rangle = 0, \qquad \chi \in \mathcal{M}_h^* . \tag{4.24}$$

It follows from (4.9) and the inverse property of $\mathcal{M}_h$ that there exists $c < \infty$ such that if $1 \le k \le r + 1$ and $w \in H_*^k(\Omega)$, then

$$\|\hat{Q}_h w - w\|_1 \le ch^{k-1} \|w\|_k . \tag{4.25}$$

The following lemmas will help us analyze T_h .

Lemma 4.1. a) Let $\vec{u} \in L^2(\Omega)^2$ and $\chi \in \mathcal{M}_h^*$. Then

$$\langle \vec{u} - \vec{Q}_h\vec{u}, \nabla\chi \rangle = 0 \ . \tag{4.26}$$

b) Let $\zeta \in H_*^1(\Omega)$ and $\vec{\chi} \in \mathcal{N}_h$. Then

$$\langle \nabla(\zeta - P_h^*\zeta), \vec{\chi} \rangle = 0 \ . \tag{4.27}$$

Proof. The proof follows easily from (4.6) and the definitions of P_h^* and $\vec{Q}_h$. Q.E.D.

Now let $\mathcal{P}_h = P_h^* \oplus \vec{Q}_h$: $H_*^1(\Omega) \times L^2(\Omega)^2 \to \mathcal{M}_h^* \times \mathcal{N}_h$.

Lemma 4.2. If $(\zeta, \vec{u}) \in D(T)$ and $(Y, \vec{v}) \in \mathcal{M}_h^* \times \mathcal{N}_h$, then

$$\begin{aligned} &B\big((\zeta,\vec{u}) - \mathcal{P}_h(\zeta,\vec{u}), (Y,\vec{v})\big) = \\ &\langle T(\zeta,\vec{u}) - T_h\mathcal{P}_h(\zeta,\vec{u}), (Y,\vec{v}) \rangle \\ &= -\langle \omega R(u - \vec{Q}_h\vec{u}), \vec{v} \rangle \end{aligned} \tag{4.28}$$

Proof. The proof follow directly from the definition of B and Lemma 4.1. Q.E.D.

Hence, it follows from Lemma 4.2 that if $Q_h = \hat{Q}_h \oplus \vec{Q}_h$, $(\zeta,\vec{u}) \in D(T)$, and $(Y,\vec{V}) \in \mathcal{M}_h^* \times \mathcal{N}_h$, then

$$(4.29) \qquad \langle Q_h T(\zeta,\vec{u}) - T_h \mathcal{P}_h(\zeta,\vec{u}),(Y,\vec{V})\rangle$$
$$= -\langle \omega R(\vec{u}-\vec{Q}_h\vec{u}),\vec{V}\rangle .$$

<u>Theorem 4.1</u>. $\delta(T_h,T) \le ch.$

<u>Proof</u>. Let $(Z,\vec{U}) \in \mathcal{M}_h^* \times \mathcal{N}_h$ and denote $T_h(Z,\vec{U})$ by $(D,\vec{E})$. We must find $(\zeta,\vec{u}) \in D(T)$ such that

$$(4.30) \qquad \|(Z,\vec{U})-(\zeta,\vec{u})\|_X + \|T_h(Z,\vec{U})-T(\zeta,\vec{u})\|_X$$
$$\le ch(\|(Z,\vec{U})\|_X + \|T_h(Z,\vec{U})\|_X) .$$

We choose $(\zeta,\vec{u})$ as follows. Let $\phi \in H_*^2(\Omega)$ be the solution to

$$(4.31) \qquad \begin{aligned} \Delta\phi &= D , && x \in \Omega , \\ \frac{\partial\phi}{\partial n} &= 0 && x \in \partial\Omega \end{aligned}$$

where $\frac{\partial}{\partial n}$ is the exterior normal derivative. (Note that $\int_\Omega D\,dx = 0$ since $D \in L^2_*(\Omega)$). Now set $\zeta = Z$, $\vec{u} = -\nabla\phi + R\nabla H\vec{U}$. It is clear that since $Z \in H_*^1(\Omega)$, $\nabla\cdot\vec{u} = -\Delta\phi = -D \in H^1(\Omega)$, $\vec{u}\cdot\vec{n} = -\frac{\partial\phi}{\partial n} = 0$ for $x \in \partial\Omega$ that $(\zeta,\vec{u}) \in D(T)$.

It follows by the definition of ϕ that $P_h^*\phi = S\vec{U}$. By elliptic regularity for (4.31),

$$\|\phi\|_2 \le c\,\|D\| \le c\,\|D\|_1 .$$

Hence, we can conclude from (4.19) that

$$\|\nabla(S\vec{u}-S\vec{U})\| = \|\nabla(\phi-P_h^*\phi)\| \le ch\|\phi\|_2 \le ch\|D\|_1 . \tag{4.32}$$

So,

$$\|(Z,\vec{U})-(\zeta,\vec{u})\|_X = \|\nabla(S\vec{u}-S\vec{U})\| \le ch\|D\|_1 \le ch\|T_h(Z,\vec{U})\|_X . \tag{4.33}$$

Denote $T(\zeta,\vec{u})$ by $(d,\dot{\vec{e}})$. It follows from (4.31) that $d = D$. Since $\mathcal{P}_h(\zeta,\vec{u}) = (Z,\vec{U})$, we obtain from (4.29)

$$\langle\vec{Q}_h\vec{e}-\vec{E},\vec{V}\rangle = \langle\omega R(\nabla\phi-\nabla P_h^*\phi),\vec{V}\rangle . \tag{4.34}$$

Thus, if we set $\vec{V} = \vec{Q}_h\vec{e} - \vec{E}$ in (4.34) and use the Cauchy-Schwarz inequality we can derive the estimate

$$\|\vec{Q}_h\vec{e}-\vec{E}\| \le c\|\nabla\phi-\nabla P_h^*\phi\| \le ch\|\phi\|_2 \le ch\|D\|_1 \le ch\|T_h(Z,\vec{U})\|_X . \tag{4.35}$$

We must now estimate

$$\|\vec{e}-\vec{Q}_h\vec{e}\| = \inf_{\vec{\chi}\in\mathcal{N}_h}\|\vec{e}-\chi\| . \tag{4.36}$$

By the definition of T ,

$$\begin{aligned}\vec{e} &= -\nabla\zeta - f\vec{u} - \omega R\vec{u}\\ &= -\nabla Z - f(-\nabla\phi + R\nabla H\vec{U})\\ &\quad - \omega(-R\nabla\phi - \nabla H\vec{U}) .\end{aligned} \tag{4.37}$$

However,

(4.38) $$-\nabla Z - fR\nabla H\vec{U} + \omega\nabla H\vec{U} \in \mathcal{N}_h$$

and

(4.39) $$\|f\nabla\phi + \omega R\nabla\phi\|_1 \leq c\,\|D\|_1 .$$

So,

$$\|\vec{e}-\vec{Q}_h\vec{e}\| = \inf_{\vec{\chi}\in\mathcal{N}_h} \|f\nabla\phi + \omega R\nabla\phi - \vec{\chi}\|$$

(4.40) $$\leq ch\,\|f\nabla\phi + \omega R\nabla\phi\|_1$$

$$\leq ch\,\|D\|_1 \leq ch\,\|T_h(Z,\vec{U})\|_X .$$

The triangle inequality, (4.35) and (4.40) yield

(4.41) $$\|T(\zeta,\vec{u}) - T_h(Z,\vec{U})\|_X = \|\vec{e}-\vec{E}\| \leq ch\,\|T_h(Z,\vec{U})\|_X .$$

Thus, (4.30) follows from (4.33) and (4.41). Q.E.D.

Now let $\lambda \in \sigma(T)$ be an isolated eigenvalue of T with finite algebraic multiplicity m. Let $E(X)$ be the generalized eigenspace corresponding to λ. It follows from Theorem 1.3 that we are interested in an estimate for $\delta(T|_{E(X)}, T_h)$.

<u>Theorem 4.2.</u> Let $\lambda \in \sigma(T)$ be an isolated eigenvalue of T and let $E(X)$ be its associated generalized eigenspace. Assume $\dim E(X) = m < \infty$, and assume that $(\zeta,\vec{u}) \in E(X)$ implies that $\zeta \in H_*^{r+1}(\Omega)$, $\vec{u} \in H^r(\Omega)^2$, $\nabla\cdot\vec{u} \in H^{r+1}(\Omega)$.

Then there exists $c < \infty$, independent of h, such that

$$\delta(T|_{E(X)}, T_h) \le ch^r . \tag{4.42}$$

Proof. Let $(\zeta,\vec{u}) \in E(X)$. Set $(Z,\vec{U}) = \mathcal{P}_h(\zeta,\vec{u})$. It follows from (4.19) and (4.22) that

$$\|(\zeta,\vec{u}) - \mathcal{P}_h(\zeta,\vec{u})\|_X \le ch^r(\|\zeta\|_{r+1} + \|\vec{u}\|_r) . \tag{4.43}$$

Also, setting $(Y,\vec{V}) = Q_h T(\zeta,\vec{u}) - T_h\mathcal{P}_h(\zeta,\vec{u})$ in (4.29) and using the Cauchy-Schwarz inequality we obtain

$$\begin{aligned} \|Q_h T(\zeta,\vec{u}) - T_h \mathcal{P}_h(\zeta,\vec{u})\|_X &= \|Q_h T(\zeta,\vec{u}) - T_h \mathcal{P}_h(\zeta,\vec{u})\| \\ &\le c\|\vec{u}-\vec{Q}_h\vec{u}\| \le ch^r\|\vec{u}\|_r . \end{aligned} \tag{4.44}$$

Finally from (4.22) and (4.24) we obtain

$$\begin{aligned} &\|T(\zeta,\vec{u}) - Q_h T(\zeta,\vec{u})\|_X \\ &\quad \le ch^r(\|\zeta\|_{r+1} + \|\vec{u}\|_r + \|\nabla\cdot\vec{u}\|_{r+1}) \end{aligned} \tag{4.45}$$

The result of the theorem now follows from (4.43), (4.44), (4.45) and the finite dimensionality of $E(X)$. Q.E.D.

Finally, we note that our results can be combined with the results in [3] to obtain optimal order estimates on the convergence of eigenvalues. Assume that the conditions of Theorem 4.2 are valid. Let $E(X^*)$ denote the generalized eigenspace associated with the eigenvalue $\bar{\lambda}$ of T^*. We note that it is well known that $\dim E(X^*) = m$ and that $\bar{\lambda}$ is an isolated eigenvalue.

Let

$$\hat{w}_h = \delta(E(X), E(X_h)) \ ,$$

$$\hat{w}_h^* = \delta(E(X^*), E(X_h^*)) \ .$$

Let α be the ascent of λ . Then it follows from Proposition 3.2 of [3] that for h_0 sufficiently small

$$\max_{i=1,\ldots,m} |\lambda - \lambda_{i,h}|^{\alpha} \le c \, \hat{w}_h \hat{w}_h^* \ ,$$

$$|\lambda - \frac{1}{m} \sum_{i=1}^{m} \lambda_{i,h}| \le c \, \hat{w}_h \hat{w}_h^* \ , \quad h \le h_0 \ .$$

Under the hypotheses of Theorem 4.2 we have proven that $\hat{w}_h \le ch^r$. Now assume, in addition, that $(\zeta, \vec{u}) \in E(X^*)$ implies that $\zeta \in H_*^{r+1}(\Omega)$, $\vec{u} \in H^r(\Omega)^2$, $\nabla \cdot \vec{u} \in H^{r+1}(\Omega)$. Then we can conclude from applying the above arguments to the adjoint problem that $\hat{w}_h^* \le ch^r$. Hence, under the above conditions we obtain the optimal order eigenvalue estimates

$$\max_{i=1,\ldots,m} |\lambda - \lambda_{i,h}|^{\alpha} \le ch^{2r} \ ,$$

$$|\lambda - \frac{1}{m} \sum_{i=1}^{m} \lambda_{i,h}| \le ch^{2r} \ , \quad h \le h_0 \ .$$

BIBLIOGRAPHY

1. J. H. Bramble and J. E. Osborn, "Rate of convergence estimates for nonselfadjoint eigenvalue approximations", Math. Comp., v. 27, 1973, pp. 525-549.
2. P. G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland Publishing Company, Amsterdam, 1978.
3. J. Descloux ,"Error bounds for an isolated eigenvalue obtained by the Galerkin method" , ZAMP, v. 30, 1979, pp. 167-176 .
4. J. Descloux, N. Nassif, J. Rappaz, "On spectral approximation, part 1. The problem of convergence" R.A.I.R.O. Numerical Analysis, v. 12, no. 2, 1978, pp. 97-112.
5. J. Descloux, N. Nassif, J. Rappaz, "On spectral approximation, part 2. Error estimates for the Galerkin method", v. 12, no. 2, 1978, pp. 113-119.
6. T. Kato, Perturbation Theory for Linear Operators, Die Grundelehren der Math. Wissenschaften, Band 132, Springer-Verlag, New York, 1966.
7. T. Kato, "Perturbation theory for nullity, deficiency, and other quantities of linear operators", J. Analyse Math., vol. 6, 1958, pp. 261-322.
8. H. Lamb, Hydrodynamics, 6th edition, Cambridge University Press, Cambridge, 1932.
9. P. D. Lax and R. Phillips, "Local boundary conditions for dissipative symmetric differential operators", Comm. Pure and Appl. Math., vol. 13, 1960, pp. 427-455.
10. J. L. Lions and E. Magenes, Problèmes aux limites non homogenes et applications, vol. 1, Dunod. Paris, 1968.
11. M. Luskin, "Convergence of a finite element method for the approximation of normal modes of the oceans", Math. Comp., v. 33, 1979, pp. 493-519.

12. W. Mills, Jr., "The resolvent stability condition for spectra convergence with application to the finite element approximation of noncompact operators", SIAM J. Numer. Anal., v. 16, 1979, pp. 695-703.

13. W. Mills, Jr., "Optimal error estimates for the finite element spectral approximation of noncompact operators", SIAM J. Numer. Anal., v. 16, 1979, pp. 704-718.

14. J. E. Osborn, "Spectral approximations for compact operators", Math. Comp., v. 29, 1975, pp. 712-725.

15. G. W. Platzman, "Normal modes of the Atlantic and Indian Oceans", Journal of Physical Oceanography, v. 5, 1975, pp. 201-221.

16. F. Riesz and B. Sz-Nagy, Functional Analysis, Frederick Ungar Publishing Co., New York, 1955.

17. J. H. Wilkinson, The Algebraic Eigenvalue Problem, Oxford University Press, 1965.

ATE
LMED
-8